







# FRANK J. SEILER RESEARCH LABORATORY

FJSRL TECHNICAL REPORT 79-0012
DECEMBER 1979

A REVIEW OF RESEARCH ON WRITTEN ADJUNCT QUESTIONS: PROGRESS AND PROSPECTS

SELECTED BODD

MAJOR JOHN N. BERMUDEZ MAJOR MARK NATAUPSKY

A

PROJECT 2303

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

AIR FORCE SYSTEMS COMMAND UNITED STATES AIR FORCE

80 6 30 078

This document was prepared by the USAF Academy Department of Behavioral Sciences and Leadership and sponsored by the Frank J. Seiler Research Laboratory, United States Air Force Academy, Colorado. The research was conducted under Project Work Unit Number 2303-F1-54, A Review of Research on Written Adjunct Questions: Progress and Prospects. Major John Bermudez was the Project Scientist in charge of the work.

When US Government drawings, specifications or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished or in any way supplied the said drawings, specifications or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

Inquiries concerning the technical content of this document should be addressed to the Department of Behavioral Sciences and Leadership, USAFA/DFBL, USAF Academy, Colorado 80840. Phone AC 303, 472-3860.

This report has been reviewed by the Commander and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

JOHN M. BERMUDEZ, Major, USAF Project Scientist

ENNETH E. SIEGENTHALER Lt COI, USAF

Director

Directorate of Chemical Sciences

WILLIAM D. SIURU, JR., Lt Col. USAF

Frank J. Seiler Research Laboratory

POHN W. WILLIAMS, JR., Coloner, USAF Professor & Head, Department of Behavioral Sciences and Leadership

Copies of this report should not be returned unless return is specified by security considerations, contractual obligations, or notice on a specific document.

Printed in the United States of America. Qualified requesters may obtain additional copies from the Defense Documentation Center. All others should apply to:

| Mational Technical Information Service 5265 Port Royal Road Springfiedl, Virginia 22161

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

	REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM				
14	FJSRL-TR-79-0012	2. GOVT ACCESSION NO. ADACISCO 77	3. RECIPIENT'S CATALOG NUMBER				
2	A BEVIEW OF RESEARCH ON WRITTEN AD PROGRESS AND PROSPECTS.	VIEW OF RESEARCH ON WRITTEN ADJUNCT QUESTIONS:					
	Z. AUTHOR(s)	M.					
	John Bermudez Major, USAF Academy Mark Nataupsky Major, USAF Academy		(6)				
	Department of Behavioral Sciences a USAF Academy Colorado 80840	and Leadership	10. PROTRAM EVEMENT, PROJECT, TASK AREA WORK UNIT NUMBERS				
	11. CONTROLLING OFFICE NAME AND ADDRESS Frank J. Seiler Research Laboratory USAF Academy, Colorado 80840	(AFSC)	December 1979				
	14. MONITORING AGENCY NAME & ADDRESS(II different	from Controlling Office)	15. SECURITY CLASS. (of this report)				
	(12	UNCLASSIFIED  15. DECLASSIFICATION/DOWNGRADING SCHEDULE					
	Approved for public release; distribution unlimited.						
	17. DISTRIBUTION STATEMENT (of the ebetrect entered	in Block 20, if different fro	m Report)				
	temporal delay motivators prose learning	cture from text naracteristics					
	ABSTRACT (Continue on reverse elde II necessory and studies on the intentional and incidinstruction/prose that was modified following written segments. The wor interpreted to suggest new lines for in the review focused on the followifactors, immediate versus delayed creerformance, and certain characteriswritten prose. Directions for future	lental acquisition by questions preciet which has been future researching variables: testing, tics of learners	n and retention of written sented before, within, or done is integrated and Studies which were included ext structure and temporal effects of motivation on in relation to instructive				
1	319920	SECURITY CLA	UNCLASSIFIED SSIFICATION OF THIS PAGE (When Data Entered)				

#### PREFACE

The authors would like to express their appreciation to all persons who offered suggestions during the preparation of this report. Special thanks are due to Dr. Raymond Kulhavy of Arizona State University for his advice and to Ms Alta Mae Miller and Ms Gladys Di Lorenzo for their outstanding typing support.

Dr. John Bermudez Dr. Mark Nataupsky

December 1979

Acces	sion For					
NTIS DDC 1	EMAS P					
Umana	fication					
By						
Distribution/						
Avel	lability Codes					
Dist.	Avail and/or special					
4						
//(						

DENNE - THE

# TABLE OF CONTENTS

				P	age
Preface		•	•	•	1
Table of Contents	•	•	•	•	11
INTRODUCTION	•	•	•	•	1
EFFECTS OF QUESTIONS		•	•		6
TEXT STRUCTURE AND TEMPORAL FACTORS	•	•	•	•	22
IMMEDIATE VERSUS DELAYED CRITERION TESTING.	•	•	•	•	32
EFFECTS OF MOTIVATION	•	•		•	35
LEARNER CHARACTERISTICS	•	•			37
SUMMARY AND DISCUSSION	•				41
SUGGESTIONS FOR RESEARCH		•		•	44
REFERENCES					46

# INTRODUCTION

This review is concerned with studies on the intentional and incidental acquisition and retention of written instructional prose that was modified by questions presented before, within, or following written segments. An attempt is made to relate these studies on the basis of variables that were common to them. These factors were either the primary or combinational objects of study, or were variables that were sufficiently controlled to enable analysis and conclusions about their effects. By examining the relatively independent effects of such variables across studies, it is possible to come to some conclusions about their individual effectiveness in improving learning from written prose.

MA DE LA

Recent interest in adjunct questions was stimulated by Rothkopf (1966), but the topic is not a new one (e.g., see Distad, 1927; Holmes, 1931; Washburne, 1929). Subsequently, Frase (1967, 1968a, b, c, d, 1970, 1971) and his associates (Frase, Patrick, & Schumer, 1970; Frase & Silberger, 1970) as well as Rothkopf and his associates (Rothkopf & Bisbicos, 1967; Rothkopf & Bloom, 1970) conducted a series of experiments on the topic. A number of other workers have been less extensively involved (e.g., Boker, 1974; Boyd, 1970, 1973; Bruning, 1968, Natkin & Stahler, 1969; Sanders, 1973; Shavelson, Berliner, Ravitch, & Loeding, 1974).

Evidently, there has been a continuing and a growing interest in research on written adjunct questions in instruction. It is equally apparent that a current and comprehensive review of the literature would have utility for assessing progress and for pointing out new directions. A few attempts at integration of the literature were of limited scope or had specific objectives. For example, Frase (1970) provided a limited summary of adjunct question research related to mathemagenic behavior, and Rothkopf (1970) described his mathemagenic behavior model. On the other hand, Bull (1973) and Ladas (1973) provided limited critical reviews. Bull (1973) focused on two research possibilities that had been relatively ignored in studies of adjunct questioning: (a) the arousal potential of questions in influencing attention, and (b) the need to measure the effect of delay on retention of questioned subjects (Ss). Ladas (1973) concentrated on a critique of the adequacy of

statistical techniques that had been used in five selected experiments and concluded that the mathemagenic hypothesis had not yet been adequately tested. In contrast, the purpose of the present review is to integrate and interpret the work that has been done and to suggest new lines for future research.

In reviewing the literature, several trends were noted which have a bearing on the content of this review as well as the interpretations and conclusions which are drawn. First, research on written adjunct questions has been a relatively specialized topic within educational psychology, although there appears to be potential for broader study on both practical and theoretical levels. A number of factors seem to have been relevant in the past: (a) the majority of the studies were carried out by comparatively few researchers; (b) only a limited number of variables have been adequately studied, and the majority were structural or orthographic in nature; (c) the experimental conditions have tended to artificially tie or separate variables to the extent that study conditions have approximated typical reading or learning environments only minimally (Hiller, 1974). The latter point is not an argument for applied research as the following discussion will indicate.

A second apparent trend is that researchers have tended to focus on one of two hypotheses: (a) the control of attention through instrumental shaping of reading behaviors. In the case of pre-questions that precede a text passage, interest has centered on selective attention.

In the case of postquestions that follow a passage, the interest has

been on forward and backward shaping resulting from progression over successive text passages. (b) The control of attention through heightened arousal induced through uncertainty about forthcoming questions. Both hypotheses have a common theoretical basis: the control of inspection behavior. It seems that other theoretical possibilities have not been given serious consideration. For example, Interference Theory, particularly proactive effects in reading, may be a fertile topic that will lead to explanations of some of the relatively consistent findings which have been produced in experiments—particularly the effect of question location in texts. The efficiency of adjunct questions may be rooted in memory factors as opposed to attentional factors.

The third trend, which is partially a result of limited theoretical interest, is that adjunct questions have been virtually overlooked as a methodological means to investigate cognitive or mediational operations in learning behavior. For example, relevant possibilities include using the effect of adjunct questions to determine how information is subjectively organized by high and low ability learners through qualitative analyses of their response errors or semantic substitutions.

It may be useful to keep these three observations on trends in mind while reading this review. Additionally, as a convenience to enable comparisons across studies, a five-category classification system will be used: (a) effects of questions, (b) text structure and temporal factors, (c) immediate versus delayed criterion testing, (d) effects of motivation, and (e) learner characteristics.

# EFFECTS OF QUESTIONS

The variables to be considered under this category include question location, knowledge of results, response mode, and type of question.

Before proceeding, it will be helpful to clarify some of the terms that will be used throughout this review.

## Definition of Terms

The positions of adjunct questions have been somewhat standardized relative to their location with respect to a written passage about which the questions are related. However, terminology varied among investigators. Questions that preceded a written passage were generally called pre-questions, whereas those that followed a passage were generally called post-questions. Questions which were presented within a related passage were variously termed embedded, inserted, interposed, or contiguous. The term embedded is, perhaps, a clearer description since some workers included both pre- and post-questions under the more general terms: inserted and interspersed, and the term contiguous lacks sufficient specificity. For the purposes of this review, pre-questions and pre-questioned groups will be referred to as PQs, and post-questions and post-questioned groups will be called SQs (subsequent question). Embedded questions will be called EQs, and No-question groups will be referred to as NQs.

# Location Effects

An apparently consistent effect of question location, found under a variety of conditions, was that SQ Ss retained more intentional or SQ-unrelated information as well as more incidental or text-related, SQ-unrelated information than PQ Ss and comparable control Ss. On the other hand, PQ Ss retained relatively little incidental information, and they retained intentional information to a lesser degree than SQ Ss (Bruning, 1968; Frase, 1967, 1968a; Rothkopf, 1966, 1970; Rothkopf & Bisbicos, 1967). Thus, in these studies, SQs seemed to produce higher overall retention than PQs, and the retention of intentional information was greater than the retention of incidental information as a consequence of SQ treatment. Higher overall retention (1.e., increased performance on tests of both intentional and incidental retention), usually has been referred to as a general facilitative effect, and higher retention of intentional information has been referred to as a specific instructive effect.

These characteristic position effects also have been replicated in comparisons of pictorial and written questions (Snowman & Cunningham, 1974), and in terms of results which can be marshalled for or against the reliability of position effects, position effects seem to be very reliable. Studies which failed to confirm these position effects include Boyd (1970, 1973; Morasky & Willcox, 1970). Boyd (1970, 1973) found no SQ facilitative effect in studies which compared treatment groups that varied mainly in position and number of questions. For all

information was less than the mean score of controls. However, the differences were not significant. In this case, failure to replicate previous findings on incidental retention could have been a function of the generally greater number of questions that were used, but this possibility seems unlikely because Boyd (1973) found a significant difference between the incidental mean for repeat-question treatments and the question-once treatment mean (p < .01). The repeat-question means (incidental, and intentional plus incidental) were above all other question-once treatment means. Morasky and Willcox (1970) found no difference between PQ and SQ groups on intentional and incidental retention. In this experiment, however, Ss were given a time-set to read as rapidly as possible whereas in the previous studies, reading was self-paced.

On the basis of these studies, then, it appears that position effects are relatively reliable, but may be subject to the influence of imposed time limits on reading, and perhaps, the number of questions asked. The effects of too little or imposed reading time seem intuitively obvious and are probably of insignificant consequence for most instructional settings since rarely are reading times imposed except im tests of reading comprehension. From the latter standpoint, there may be some payoff in studying position effects and time. Such applicatins have not been attempted by any of the studies reported in this review. The effect of position and number of questions across equal size segments of text has not been studied directly. Future studies could vary

the number of questions to determine the optimum number of PQs or SQs that will produce optimum overall retention. It could also be that response-induced output interference impedes intentional and incidental retention in differential ways.

# Effect of Knowledge of Results (KR)

Only a few studies provided KR (e.g., Bruning, 1968; Frase, 1967; Rothkopf, 1966; Shavelson et al., 1974). Bruning (1968) provided KR on SQ; however, the design of the study did not allow an evaluation of KR compared to groups not given KR. Frase (1967) found a significant interaction between question position and whether or not a correct answer was available (p < .05). When KR was provided, there was little difference between PQ and SQ groups on intentional retention; however, when KR was not provided, SQ Ss performed significantly better. This relation did not hold up with incidental retention.

Frase's (1967) results with KR were not consistent with those obtained by Rothkopf (1966). In Rothkopf's (1966) study, KR was provided to one of two PQ groups (the SBA group) and one of two SQ groups (the SAA group) on intentional information. As might be expected, both the SBA group and the SAA group responded correctly on a test of intentional information more often (p < .01) than groups who did not receive KR (the SA and SB groups). The SBA group obtained a mean score approximately 19 percent higher than the SB group, and the SAA group obtained a mean score approximately 31 percent higher than the SA group. This relation did not hold up on a test of incidental information. The SBA group

575 (144) A. C.

obtained a mean score 15 percent higher than the SB group, while the SAA group obtained a mean score approximately 7 percent lower than the SBA group.

In the Frase (1967) study, then, KR was not facilitative in learning intentional information, whereas in the Rothkopf (1966) study, it was. With incidental information, KR made no difference whereas in the Rothkopf study, it had a facilitative effect for PQ groups and an inhibitory effect for SQ groups. Although both studies provided effective control for accidental exposure of the correct answers, the Rothkopf (1966) study was probably the more effectively controlled since there was absolutely no opportunity for pre-searching answers. Insufficient control of pre-search availability leads to confused results. For example, Shavelson et al. (1974) provided KR to four experimental groups. The control group simply read the text. Except for an SQ group that received higher order (application) questions, the control group performed as well or better than the experimental groups. The design of the materials, however, allowed PQ groups to pre-search answers since KR was presented to them at the end of each text section. SQ groups, on the other hand, received KR on a following page.

What other factors could account for such different results in the Frase and Rothkopf studies? Aside from content, the studies differed primarily in response mode, length of text, and pacing or frequency of questions. Frase (1967) used multiple-choice questions, and Rothkopf (1966) used constructed responses. The text used by Rothkopf (1966) was

more than double in length than the text used by Frase (1967). The latter factor in combination with variable pacing in the Frase study could have produced the effect since Frase (1967) found that scores on incidental retention improved with passage length. Nevertheless, these conjectures should be put to experimental test. It seems clear that more work is needed to determine the differential role of KR on intentional versus incidental learning aided by adjunct questions.

# Effect of Response Mode

Approximately one-half of the studies reviewed have employed multiple-choice questions as adjuncts and for the criterion tests. The number of response alternatives varied from three choices to five choices. Almost all these studies essentially replicated the characteristic location effects of PQs and SQs. An exception was the Morasky and Willcox (1970) study which employed three alternative choice questions. The balance of the remaining one-half of the studies consisted almost entirely of constructed responses of the one to two missing terms variety. Natkin and Stahler (1969) used a short answer mode, and essentially replicated the earlier position effects as did almost all other studies which used constructed responses. An exception was Boyd (1973) who found no general facilitative effect for SQs. However, Boyd's (1973) study differed from the other studies in the number of questions asked per segment of text.

Frase (1968) varied response modes in one design and found no significant differences between groups using different modes. He compared

multiple-choice questions consisting of five alternatives and constructed responses which consisted of interrogatives that required Ss to supply a critical word or words. A free-recall mode was employed by Frase (1971) involving free recall of facts and inferences immediately after reading. Ss had to write all they could remember about text passages, including facts and any valid inferences they could generate. The results of this study are discussed in the next section under question type, since this study did not compare performance of PQs versus SQs.

The effect of response mode appears to make little difference on the characteristic effects of PQs and SQs in terms of multiple-choice questions versus incomplete sentences with one and two missing terms. These conclusions can only be considered as tentative since it is possible that number of alternatives in multiple-choice questions may differentially affect results, but that is unlikely. It is clear, however, that there is too little data on short-answer and free-recall response modes to make any conclusions about these forms of responses and their interaction with position effects.

#### Effect of Type of Question

A potentially fruitful area for research is in examining the effects of type of question. Little work has been done from either a practical, pedagogical standpoint or with a view to developing some theoretical bases for cognitive studies. Some difficulties that occur in comparing experiments on the basis of type of question used for adjunct purposes or for criterion testing include: (a) insufficient information in the

research reports regarding rules of formation of the questions or the theoretical framework involved, (b) the relative infrequency of studies that specifically manipulated type of questions as a variable, and (c) variability in terminology used to describe the items. Frase (1970) agreed that development of a precise taxonomy for the experimental manipulation of questions is sorely needed. How, then, may we compare experiments on some common basis? For the purposes of this review, question types will be categorized under two general kinds: those that required reproduction of information and those that required Ss to make near or remote associations.

# Reproduction Types

Reproduction will refer to the demonstration of knowledge in essentially the same semantic or orthographic form in which it was presented. Studies which varied reproduction questions were conducted by Frase (1968d) and Rothkopf and Bisbicos (1967). Studies which did not vary reproduction type questions but which used this form to investigate the effects of other variables include: Boker (1974), Boyd (1970), Frase (1968b), Frase and Silbergen, (1970), Morasky and Willcox (1970), Natkin and Stahler (1969), and Rothkopf and Bloom (1970).

Frase (1968c) presented three types of reproduction PQ that related to a 36-word highly structured paragraph and allowed three groups of Ss 20 seconds reading time. The questions: general, specific, and comparative, varied in the amount of information to which they related (i.e.,

number of words necessary to answer the PQ). The results of an immediate criterion test indicated that both intentional and incidental retention was a function of the type of PQ. The least learning occurred with general questions, whereas specific questions produced the most learning. Rothkopf and Bisbicos (1967) studied the specific instructive effects i.e. intentional retention) of question type on a criterion test administered to high school students immediately after exposure to a 36-page text. Two questions preceded or followed every three pages. The categories of questions were: (a) common phrases consisting of common, non-technical words (C); (b) technical and scientific words (T); (c) measurement terms dealing with size, distance, dates (M); and (d) names such as proper geographical and personal names (N). An NQ control and six treatment groups which varied as to question position (PQ or SQ) and pair-type of questions (i.e., N-M, C-T, or mixed (MX)) were compared. A selective effect of type of question was inferred from the higher performance on the criterion test of type of response associated with type of PQ or SQs that had been presented in the text. The MX and N-M SQ groups scored the largest average number of correct responses on M and N items. In a like manner, the MX and CT groups produced a greater number of correct responses in the C-T phrase category than other treatments. Thus, Frase's (1968c) and Rothkopf and Bisbicos' (1967) data would lead us to the tentative conclusion that specific questions are more facilitative than general questions and that specific questions lead to type-specific reproduction of information. It should be recalled, however, that Morasky and Willcox (1970) used common word

questions similar to Rothkopf and Bisbicos (1967) and found no significant differences between PQ and SQ groups.

# Near and Remote Association Types

Near and remote association will refer to the demonstration of correct responses as a consequence of the application, analysis, or synthesis (Bloom, 1956) of near and remote facts or inferences which can be obtained or abstracted from a text.

Frase (1971) studied the effects of inference and fact questions on free recall and found no statistical differences between inference and fact groups. On the other hand, free recall of both facts and inferences revealed that inferential PQs produced higher recall than factual PQs (p < .001). Factual questions produced 9.8 percent recall of incidental facts whereas inferential questions produced 35.4 percent recall of facts. The mean recall of the inference group for recall of incidental inferences was 17.3 percent and for the fact group, it was 1.4 percent. Evidently, inferential adjunct questions are poor for specific or intentional instruction of inferences, but produce recall of incidental facts as much as factual questions do.

Watts and Anderson (1971) investigated the effects of SQs that required Ss to transfer what they had read to a new situation. There were six treatment groups, each of which received one type of SQ which followed each of five 450-word passages. The criterion test consisted of 25 questions composed of five of each of the five following types: Groups RE, and RE, (repeated examples) received SQs that were later

repeated exactly as correct alternatives to a multiple-choice question in the criterion test. Groups  $A_1$  and  $A_2$  received the same questions as the RE groups except that the correct multiple-choice alternatives described an example of a concept or a principle that was different from any that were employed in the passage. Group N received questions in which correct alternatives were names of persons associated with a principle or concept. Group C was a reading-only control group. Both groups who answered application questions during instruction performed significantly better overall than the RE, N, or C groups (p < .05). There was considerable variation in group performance on the different types of questions (p < .01). The application question groups demonstrated a significantly higher overall performance than the remaining groups (p <.01). Performance by all groups on the N questions was low; however, the RE, group was significantly below all other groups on the N questions. This group repeated the example that occurred in the opening paragraph of the passage. There were no group differences on repeated example questions. A possible criticism of this study is that the results could have been a function of difficulty level of questions. An interesting finding, however, is the reversed performance on the SQs (i.e., reverse group rankings) for time spent on SQs versus the rankings on the criterion test. The reverse rankings indicated, perhaps, that high performance on the criterion test was partially a function of time spent on the SQs in the text and also that high performance on SQs did not insure a similar level of performance on the criterion test.

Unfortunately, almost all the studies on adjunct questions did not

THE THE TOTAL

examine time spent on PQs and SQs, and obviously more work is needed in this area, especially with respect to type of question, time on questions, and differential criterion performances.

A second study which compared reproduction and near and remote association questions, was conducted by Shavelson et al. (1974). These investigators presented a 1,525-word text and varied type of question: higher-order (application and analysis) and lower-order (factual reproduction) and location (PQ and SQ). Four experimental groups had either one low or one higher-order multiple-choice PQ or SQ presented along with each of eight sections of the text. Each group was thus given a total of eight questions. Since in this study, the assumption of homogeneity of variance was violated, the results must be cautiously interpreted. The SQ higher-order question group scored substantially higher than the PQ lower-order group on both an immediate and a delayed criterion test for intentional and incidental retention. No other reliable differences were found. Despite the no difference findings, contrasts between types and positions of questions were performed. The group means reflected a greater facilitative effect for the higher-order SQ than for other comparisons. Unexplainably, however, the NQ control group performed as well or better than the experimental groups. This study also included the administration of five ability measures to determine the effect of individual differences. This aspect of the study will be discussed under the section on learner characteristics.

Aside from the obvious differences which distinguish this study from

others on type of question, Shavelson et al. (1974) apparently presented PQs and SQs along with the text. This procedure, while it approaches the typical reading task environment, may have tended to wash out some of the effects between groups. This is a reasonable interpretation since most of the researchers have presented text sections and adjunct questions on separate pages, and because group mean differences in retention have been relatively small under these atypical reading conditions.

What conclusions may we draw regarding types of questions? The results apparently are ambiguous as well as inconsistent, and it is probably premature to make any definitive statements beyond the fact that some types of specific questions do seem to produce selective differences; however, it is not very clear how the differences interact with location of a question in the text, nor is enough known about the relation of time spent on adjunct questions by type of question and how these variables affect criterion performance.

## TEXT STRUCTURE AND TEMPORAL FACTORS

The variables that will be discussed under this section include frequency and pacing of adjunct questions and reading time.

#### Frequency and Pacing

The effect of frequency and pacing of questions in text was varied experimentally by Frase (1967, 1968a, 1968b) and Frase et al. (1970).

In reviewing these results, it should be remembered that conclusions about intervals between questions may be influenced by the units of measure used in various studies. In some cases, lines of type were used to vary the intervals (e.g., Frase, 1967). Other studies based variations on number of sentences without restrictions on the length of sentences (e.g., Frase, 1968a). Most studies, however, referred to the number of words per section, paragraph, or page of a passage.

Number of typed lines between questions. Frase (1967) varied intervals with lines of type between questions and found that the effect of question pacing tended to be different for the retention of intentional and incidental information. In contrast to a curvilinear relation for intentional retention, across 10, 20, or 40 lines of type, there was a slight gradual improvement in scores on incidental retention with the longer passages. A passage of 20 lines was optimal, however, for both intentional and incidental retention.

Frase, Patrick, and Schumer (1970) examined the combined effect of incentive and question frequency in a 2,000-word text used in the Frase (1967) study. Since the report of the study included no details about passage interval in terms of sentences or lines, it is necessary to make some assumptions. There were 20 paragraphs presumably averaging 100 words each. Assuming 10 words per line, each paragraph consisted of 10 lines. Frequency, then, was probably defined as one question before or after every paragraph of 10 lines and infrequency was defined as five questions before or after every five paragraphs of 50 lines. PQ groups scored .66 questions above controls on intentional retention test items,

but they averaged 2.0 questions below them on incidental retention items. When questions were infrequent, Ss in both PQ and SQ groups performed above controls on both intentional and incidental retention.

The Rothkopf and Bisbicos (1967) study can also be looked at for effects of infrequent pacing. In this study, a 9,000-word text was divided into 36 pages of 250 words each. Two questions were presented before or after every three pages. The results showed that treatments involving SQs resulted in better overall retention than PQ treatments and the NQ control treatment. There was no difference between the PQ and control group.

Number of sentences between questions. Frase (1968a) presented Ss with a 2,000-word passage and questions before or after every 10, 20, 40, or 50 sentences. The same questions were used under all four conditions, but the pacing of the questions was different for each group. Thus, for the different groups, one question was presented after each 10 sentences, two questions after every 20 sentences, four questions after every 40 sentences, or five questions after every 50 sentences. In this study, there was no difference between PQ and SQ groups in the 50-sentence condition. However, retention of incidental information was substantially lower for PQ Ss when the questions occurred every 10 sentences, whereas for the SQ Ss, retention of both intentional and incidental information was highest. When a question occurred after every 10 sentences, the SQ groups scored about 40 percent higher on overall retention than the PQ groups.

Number of words between questions. Frase (1968b) presented a

2.000-word passage divided into 20 paragraphs of 100 words each and varied the frequency of questions (after every 10 and 20 sentences) and blocks of paragraphs for PQ and SQ groups. There were four blocks of five paragraphs each. The dependent measure was the number of items correct for intentional and incidental retention of five items from each of the four-paragraph blocks. The results showed that the SQ group performed at a higher level than the PQ group on both intentional and incidental information (p<.005). The effect of frequency was different for PQ and SQ groups. Although there was no significant effect for frequency alone, the interaction of frequency with question location was significant at the .05 level. The mean scores on intentional and incidental retention were higher when PQs were presented before every other paragraph than when presented before every paragraph. However, both intentional and incidental score means were higher when questions were presented after every paragraph. Increased frequency, therefore, favored the SQ group. The effect of blocks was examined in the SQ group to determine if a learning-to-learn effect had occurred. The main effect for blocks was significant at the .05 level, but the interaction of question group (PQ or SQ) and blocks was not significant. A trend analysis of the block effect revealed a significant cubic component (p < .025).

Rothkopf and Bloom (1970) divided a 16,200-word text on earth sciences into 108 pages with approximately 100 to 200 words per page.

Thirty-five mm photographic slides were prepared for each page. Beginning with the 25th slide, a written intentional retention SQ was placed

in the slide sequence after every sixth slide for one group of Ss whereas identical questions were presented orally to a second group of Ss. Although the purpose of the experiment was to explore the influence of orally directed questions (social contact), the extreme length of the text in comparison to most studies and the superior performance of written SQ groups over the NQ control group is especially noteworthy.

Natkin and Stahler (1969) presented a 1,500-word passage divided into 25 paragraphs of 100 words each. Each paragraph was followed by an SQ. SQ Ss exceeded NQ Ss on an incidental retention test.

In contrast to most experiments, Boyd (1973) failed to find an interaction between two levels of frequency (one and two paragraphs), using 100 words per paragraphs, and PQ and SQ groups. In a study involving only PQ groups, Peeck (1970) gave a 3,000-word text presumably distributed over six pages, each containing 500 words. PQ groups outperformed reading control groups on intentional retention, but there were no differences between PQ groups and an extended-reading control group which was given an extra amount of reading time equal to the time for PQ Ss to read and answer PQs. Interestingly, a reading-only control group performed significantly lower than all other groups.

The results of the majority of studies on question interval effects may be summarized in capsule form as follows: Approximately 20 lines of type or 250 words may be optimal for intentional and incidental retention of specific factual information, and longer passages may improve incidental learning somewhat. While mean performance score differences are not very great between PQ and SQ groups, differences between groups

given adjunct questions and those not given questions are much larger and hold up over extremely long passages. When questions are massed versus distributed over text, there is no difference between PQ and SQ groups on long passages (e.g., over 50 lines or sentences); however, frequent SQs seem to improve retention substantially more than frequent PQs. The two latter studies by Boyd (1973) and Peeck (1970) deviate from most of the other studies in ways which may account for the failure to replicate question interval effects. The Boyd (1973) study used more adjunct questions per portion of text than all previous studies, and the Peeck (1970) study included a fixed reading time and 15 massed PQs presented before reading. Thus, the number of questions presented to Ss is no doubt an important variable of adjunct questioning. It is very likely that by increasing the number of questions beyond an optimum number, memory becomes overloaded and impedes processing of information from the text (PQ) or from memory (SQ). This may be a problem which can be readily explored in a discrete study.

#### Reading Time

Reading time or the time an <u>S</u> spends on all or a portion of a text has been used as an important indicator of control of attention. Conversely, it is apparent that reading time has not been adequately nor sufficiently investigated to confirm theoretical explanations concerning the operant effects of SQs. Currently accepted explanations about PQs (e.g., see Frase, 1970) are probably warranted in light of the conclusions which have been drawn from research on pre-search availability

(e.g., see Anderson & Faust, 1967). However, the role of SQs seem much less clearer.

Natkin and Stahler (1969) gave Ss 15 minutes to read a 2,500-word biology text and found non-significant differences between SQ and NQ groups. Morasky and Willcox (1970) instructed Ss to read as quickly as possible a 2,000-word text and found that a PQ group mean completion time was significantly less than an SQ group. Rate of comprehension scores were computed for each S using the formula: Rate of Comprehension = percent of test items correct X completion time. The PQ group rate was significantly greater than that of the SQ group on: intentional test scores (p < .001), incidental scores (p < .01), and total test scores (p < .001). In a second follow-up study in which Ss were timed individually to the nearest second after each one-paragraph page was turned, Ss were told they were being given a speed test. No significant differences were found between PQ and SQ groups on total completion time defined as time taken for both reading and answering questions. The PQ paragraph completion time was significantly less than the SQ time (p < .01). The SQ group required significantly less total time to complete intentional questions than the PQ group (p < .005). Although the PQ group required less time on all 21 paragraphs, the differences were not significant until after over one-half of the paragraphs had been completed.

Watts and Anderson (1971) found that reading time decreased significantly over five 450-word passages (p<.01) but found no significant differences between groups in the mean total time spent on all passages.

Note that the

Rothkopf and Bisbicos (1967) had <u>Ss</u> record the time they started and finished each page. There was no reliable difference in mean reading time per page among six treatment groups and one control group. However, the reading time per page declined as a function of the number of pages read, and the average inspection time per page was significantly shorter for the second one-half of the text than for the first one-half (p < .001). Rothkopf and Bisbicos (1967) suggested that these results were indicative of an extinction-like process which progressively weakened inspection behavior during the course of study. Also, the effect of SQs tended to be more marked for retention of the second half of the text than for the first half so that SQs appeared to be more effective in the second one-half of the text than in the first half.

Interestingly, these same data could be explained by a build-up and release from proactive inhibition (e.g., see Wickens, 1970), yet interference theory was not considered in any of these studies. Nor has sufficient attention been given to the time an S spends on adjunct questions in relation to the time spent on portions of the text. Trend analyses and differential ratios of changes over time could prove to be invaluable from both practical and theoretical standpoints.

# IMMEDIATE VERSUS DELAYED CRITERION TESTING

With few exceptions, criterion tests were administered to Ss either immediately after reading a text or after a brief filler-task designed to minimize the effects of recency in recall. Several studies did not include a control for recency (e.g., Frase, 1968b, 1971; Frase and

Silberger, 1970; Rothkopf & Bisbicos, 1967; Watts & Anderson, 1971; Peeck, 1970). The exceptions compared the effect of immediate versus delayed testing (usually seven days). Natkin and Stahler (1969) compared SQ and NQ groups and found that the SQ  $\underline{S}$ s performed significantly better (p < .01) on a test of incidental retention. However, this result should be cautiously interpreted since Natkin and Stahler (1969) did not control for transfer between intentional and incidental questions as previously outlined by Rothkopf (1966) and as later recommended by Ladas (1973).

Other studies which either employed no transfer control or controlled transfer by logical classification of adjunct questions include: Bruning (1968), Boker (1974), and Peeck (1970). However, Bruning (1968) performed a post hoc analysis which indicated that transfer effects had not influenced the results, and Boker (1974) controlled for question difficulty level—a control not used by most workers.

Peeck (1970), on both an immediate and a delayed test of incidental retention, found that an NQ-extended reading control group, which was given about five minutes extra reading time, scored significantly higher (p < .05) than PQ groups which had been questioned over intentional information, even though total experimental times were equalized. Boker (1974) presented a 2,500-word passage divided into 10 sections of 250 words each. Two questions on intentional information were inserted before or after each related section. A control group read the text without questions (NQ). The test-time effect was significant (p < .001) as expected. The retention means decreased about 10 percent over the

seven days, and the same trend was found for both the immediate and delayed measures. Comparison of both PQ and SQ retention means on intentional information revealed that both the PQ and SQ means were significantly larger than the control group mean (p < .05), but they were not significantly different from each other. Analysis of retention means on incidental information, collapsed over time, indicated that the PQ and NQ control means were not significantly different from each other but both were significantly lower than the SQ mean (p < .05).

Thus, the specific instructive effects and general facilitative effects were confirmed as holding over time in this study. However, since studies which compare immediate and delayed retention resulting from adjunct questions are a relative rarity, the results should be accepted on a tentative basis (e.g., see Hiller, 1974).

#### EFFECTS OF MOTIVATION

Motivation has been investigated mainly in terms of monetary incentive, and the arousal-producing nature of adjunct questions.

#### Monetary Incentive

Frase, Patrick & Schumer (1970) varied the influence of monetary incentive for PQ and SQ groups as well as a control group who received 0¢, 3¢, and 10¢ for each correct criterion response. SQ Ss scored significantly higher than PQ Ss, and although the 3¢ and 10¢ group scores did not differ from each other, both groups differed significantly

from the 0¢ group. At the 0¢ level, the SQ group differed from the PQ group. At the 3¢ level, both the control and SQ groups differed from the PQ group, while no comparisons were significant at the 10¢ level.

Frase (1971), in a study of incentive and type of adjunct questions, presented three short passages, each consisting of four experimental text sentences of different logical classes embedded in other information. The logical classes were of the form: As are Bs, Bs are Cs, Cs are Ds, and Ds are Es. Two PQs appeared above each passage in the form of a sentence. Ss had to determine whether the conclusion of the sentence could be drawn from the passage below it. Ss then circled one of two response options: "valid" or invalid." All the Ss were paid volunteers. Incentive was varied by telling one-half of the sample they could earn additional money, and the other one-half that they were not in the paid group. Timing of incentive was also varied by informing Ss before or after reading the text (incentive information). Incentive information, provided before reading, produced more correct solutions on the PQ than after reading (p < .05), and pay produced more correct solutions on the PQs than no pay (p < .01). Performance on recall of incidental information was also higher when incentive information was presented before reading (p <.005). However, overall, the incentive level between fact and inference groups was not significant.

## Arousal

Natkin and Stahler (1969) tested the hypothesis that questions function as arousal stimuli by presenting two 2,500-word passages successively to four groups of Ss. Group A received SQs on Text #1 and SQs on Text #2 (high pre-exposure to questions). Group B received no questions on Text #1 and SQs on Text #2 (low pre-exposure to questions). Group C received SQs on Text #1 and NQ on Text #2. Group D got NQ on either text. Under conditions of high pre-exposure to questions, there was a typical decline in retention from immediate to delayed testing. When questions were introduced with no pre-exposure, they resulted in a marked increase in delayed performance. Although the results substantiated Natkin and Stahler's (1969) predictions, they were also consistent with the characteristic findings concerning facilitative effects of SQs on incidental learning.

As may be expected, incentive helps performance, and since mean differences between groups, although statistically significant, are not very great, incentive may overcome location effects. There is no doubt that arousal heightens attention, and attention improves recall, but thus far the work on arousal has not shown whether SQs produce more arousal that PQs, though it is likely that they do.

#### LEARNER CHARACTERISTICS

Learner characteristics or individual differences have only recently been considered by investigators of adjunct questions. This new development is no doubt a reflection of growing interest in this line of research, but also a sign of greater sophistication in studying the relevant variables, and a sign of greater concern with practical applications.

Shavelson et al. (1974) gave five ability measures to his Ss and found that only a vocabulary test, which measured comprehension interacted with treatment. There was a slight negative correlation between scores on the comprehension measure and the criterion measure for Ss in the higher-order SQ groups, indicating, perhaps, that higher-order questions interfered with learning of Ss who already had a high level of verbal comprehension.

Sanders (1973) investigated the interaction of individual differences and adjunct questions by blocking on grade-point average, and contrary to finding characteristic location effects, he found no significant differences between PQ and SQ groups on an immediate test. However, he found that high ability SQ Ss outperformed all other Ss in both the PQ and SQ conditions. No difference was found between high and low ability Ss in NQ conditions. Apparently, when adjunct questions were present, high ability Ss took advantage of them. Sanders (1973) suggested that his failure to replicate position effects may indicate that the sensitivity of type of written materials and its relation to individual ability is important; however, since he used the same materials that Frase (1967) used, and since position effects have held up over a variety of subject matters, it appears more likely that ability is the more relevant variable.

Swenson and Kulhavy (1973) attempted to replicate and extend position effects to a grade school population of fifth and sixth graders. They

+

tested the hypothesis that adjunct questions act on the type of encoding in which readers engage—that is, phonological versus semantic encoding (see Anderson, 1972). The pupils were presented with twenty 66—word paragraphs with the number of PQs of SQs, inserted before or after one, five, 10, or 20 paragraphs, corresponding to the number of paragraphs. All paragraphs had been normed for readability and semantic similarity—a procedure not usually employed in this line of research. As in most studies, transfer control between intentional and incidental information was done logically.

Two types of questions were used: an original form or a lexical paraphrase having no words in common with the original form except for articles and terminal response. Immediate and delayed criterion measures included both intentional and incidental items. The results showed that SQ groups outperformed PQ groups (p < .05), but no difference was found on lexical and paraphrased items. In addition, SQs failed to differentially facilitate retention of incidental information.

An interesting finding was that overall learning and retention was optimized over time most for Ss who were given questions at intervals of five paragraphs (330 words or approximately 30 lines). In contrast, Frase (1968b) using an adult population found that overall retention was optimized at about 20 lines of type or 250-word intervals. This difference between adults and children is consistent with Swenson and Kulhavy's (1974) interpretation that longer reading segments may serve to assist children in acquisition of passage theme, thus improving retention.

Adults, being more sophisticated verbal learners, are able to extract

greater meaning from shorter segments.

The Shavelson et al. (1974) and Sanders (1973) results essentially indicate that individual ability is an important variable in determining the effectiveness of adjunct questions. However, to further delineate specific effects of adjunct questions, it may be wise to separate the effects of ability and achievement motivation. Additionally, adjunct question research should be extended to lower grade levels since almost all the research has been done on undergraduates.

#### SUMMARY AND DISCUSSION

In this review, we have examined the effects of adjunct questions, text structure and temporal factors, immediate versus delayed criterion testing, the effects of motivation on performance and certain characteristics of learners in relation to instructive written prose. Within this context, it was suggested that location effects are reliable, but may be sensitive to imposed time limits on reading and the number of adjunct questions asked per segment of text. The number of questions asked per segment over succeeding segments is probably an important variable in keeping learning and retention high. In studies on frequency, questions have been either massed or distributed in text, but the effect of variation in number on serial position of segments has not been studied.

A weakness of most studies is inattention to the temporal aspects of reading text compared to time spent on adjunct questions. Another problem is that theoretical explanations of the mechanisms underlying the effects of SQs have been restricted to inferences drawn from S time spent

on reading text and magnitude of criterion scores. Differences between PQ and SQ groups have been explained in terms of inspection time which is thought to be shaped instrumentally. Alternative explanations which can be derived from interference theory or memory load considerations have not been tested. Such alternative hypothesis may be more relevent to SQs than to PQs.

Only a few studies provided KR, and the effects of this variable are not clear. For example, the provision of KR is almost like allowing Ss to turn back to a previous page, but probably reduces general facilitative effects. Also, the differential role of KR on intentional versus incidental learning and retention has not been sufficiently worked out in the context of adjunct questioning.

The form of response mode appears to make little difference on retention, but there is too little data on short-answer and free-recall modes for a definitive conclusion about this variable in connection with adjunct aids.

Comparisons of the effects of types of questions suffer from lack of a precise taxonomy for experimental purposes; however, past results seem to indicate that specific, factual questions are more facilitative than general or inferential questions for providing learning of both facts and inferences, intentionally or incidentally. This effect is probably a function of the succinct conveyance of meaning to a reader and the interaction of meaning with idiosyncratic methods of forming inferences.

A majority of criterion tests were administered immediately. Frequently in these studies, no controls were imposed to minimize the effect

of recency. In addition, few workers used controls to insure the absence of transfer between intentional and incidental information. A few studies have confirmed that specific instructive effects maintain over time, but additional data should be accumulated on this topic since delayed measures were rarely used.

Studies involving motivation and individual differences have shown that increased incentive allows Ss to overcome differences induced by question location, and there is some indication that Ss with high scholastic ability or high verbal skills benefit less from adjunct questions than low ability Ss, but very little is known about the interaction of treatments and aptitudes.

## SUGGESTIONS FOR RESEARCH

Future research on adjunct questions should probably concentrate on the nature and conditions of SQs in facilitating both intentional and incidental learning and retention. Studies which involve variation in number of SQs on serial position of text segments could lead to theoretical advances as well as practical pedagogical applications. In this connection, accumulation of data on reading times per segment and per related set of questions as well as criterion scores per segment would seem potentially useful. It would also be interesting to determine how KR on intentional material differentially affects incidental learning. Future studies should also attempt to control for recency, and transfer. If logical methods are used to control for transfer, post hoc analyses should be performed to confirm the absence of transfer. In addition,

reference : St

delayed measures should be taken whenever feasible. The effect of monetary incentive seems relatively clear, but it is difficult to see any utility for exploring this variable further. Perhaps blocking on grade-point average and measures of achievement motivation is more relevant. Other relevant differences, for example, of scholastic ability, verbal comprehension, and attitudes towards adjunct aids appear to be fruitful areas of research. Another possibility might be to block on measures of short-term memory for sentences or other verbal units to determine the differential effectiveness of adjunct questions.

## REFERENCES

- Anderson, R. C. How to construct achievement tests to assess comprehension. Review of Educational Research, 1972, 42, 145-170.
- Anderson, R. C. & Faust, G. W. The effects of strong formal prompts in programmed instruction.

  <u>Journal</u>, 1967, <u>4</u>, 345-352.
- Bloom, B. (Ed.) <u>Taxonomy of educational objectives</u>, <u>handbook I</u>; <u>Cognitive domain</u>. New York: Longmans, Green & Co., 1956.
- Boker, J. R. Immediate and delayed retention effects of interspersing questions in written instructional passages. <u>Journal</u> of <u>Educational Psychology</u>, 1974, 66, 96-98.
- Boyd, W. M. The effect of test-like events on attention and retention in prose learning. Paper presented at the meeting of the American Educational Research Association, Minneapolis, March 1970.
- Boyd, W. M. Repeating questions in prose learning. <u>Journal of Educational Psychology</u>, 1973, 64, 31-38.
- Bruning, R. H. Effects of review and test-like events within the learning of prose material. <u>Journal of Educational Psychology</u>, 1968, 59, 16-19.
- Bull, S. G. The role of questions in maintaining attention to textual material. Review of Educational Research, 1973, 43, 83-87.
- Distad, H. W. A study of the reading performance of pupils under different conditions on different types of materials. <u>Journal of Educational Psychology</u>, 1927, 18, 247-258.
- Frase, L. T. Learning from prose material: length of passage, knowledge of results, and position of questions. <u>Journal of Educational Psychology</u>, 1967, 58, 266-272.
- Frase, L. T. Effect of question location, pacing, and mode upon retention of prose material. <u>Journal of Educational Psychology</u>, 1968, 59, 244-249. (a)
- Frase, L. T. Some data concerning the mathemagenic hypothesis.

  American Educational Research Journal, 1968, 5, 181-189. (b)

Translation Design

- Frase, L. T. Some unpredicted effects of different questions upon learning from connected discourse. Journal of Educational Psychology, 1968, 59, 197-201. (c)
- Frase, L. T. Questions as aids to reading: Some research and theory. American Educational Research Journal, 1968, 5, 319-332. (d)
- Frase, L. T. Boundary conditions for mathemagenic behaviors.

  Review of Educational Research, 1970, 40, 337-347.
- Frase, L. T. Effect of incentive variables and type of adjunct question upon text learning. <u>Journal of Educational Psychology</u>, 1971, 62, 371-375.
- Frase, L. T., Patrick, E. M., & Schumer, H. Effect of question position and frequency on learning from text under different levels of incentive. Journal of Educational Psychology, 1970, 61, 52-56.
- Frase, L. T. & Silberger, F. Some adaptive consequences of searching for information in a text. American Educational Research Journal, 1970, 7, 553-560.
- Hiller, J. H. Learning from prose text: Effects of readibility level, inserted question difficulty and individual differences.

  <u>Journal of Educational Psychology</u>, 1974, 66, 202-211.
- Holmes, E. Reading guided by questions versus careful reading and re-reading. The School Review, 1931, 39, 361-370.
- Ladas, H. The mathemagenic effects of factual review questions on the learning of incidental information: A critical review.

  Review of Educational Research, 1973, 43, 71-82.
- Morasky, R. L. & Willcox, H. H. Time required to process information as a function of question placement. American Educational Research Journal, 1970, 7, 561-567.
- Natkin, G. & Stahler, E. The effects of adjunct questions on shortand long-term recall of prose materials. American Educational Research Journal, 1969, 6, 425-432.
- Peeck, J. Effect of prequestions on delayed retention of prose material. <u>Journal of Educational Psychology</u>, 1970, 61, 241-246.
- Rothkopf, E. Z. Learning from written materials: An exploration of the control of inspection behavior by test-like events. American Educational Research Journal, 1966, 3, 241-249.

- Rothkopf, E. Z. The concept of mathemagenic activities. Review of Educational Research, 1970, 40, 325-333.
- Rothkopf, E. Z. & Bisbicos, E. E. Selective facilitative effects of interspersed questions on learning from written materials.

  Journal of Educational Psychology, 1967, 58, 56-61.
- Rothkopf, E. Z. & Bloom, R. D. Effects of interpersonal interaction on the instructional value of adjunct questions in learning from written material. <u>Journal of Educational Psychology</u>, 1970, 61, 417-422.
- Sanders, J. R. Retention effects of adjunct questions in written and aural discourse. <u>Journal of Educational Psychology</u>, 1973, 65, 181-186.
- Shavelson, R. J., Berliner, D. C., Ravitch, M. M., & Loeding, D. Effects of position and type of question on learning from prose material: Interaction of treatments with individual differences. Journal of Educational Psychology, 1974, 65, 40-48.
- Snowman, J. & Cunningham, D. J. A comparison of pictorial and written adjunct aids in learning from text. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1974.
- Swenson, I. & Kulhavy, R. W. Adjunct questions and the comprehension of prose by children. <u>Journal of Educational Psychology</u>, 1973, 66, 212-215.
- Washburne, J. N. The use of questions in social science material.

  <u>Journal of Educational Psychology</u>, 1929, 20, 321-359.
- Watts, G. H. & Anderson, R. C. Effects of three types of inserted questions on learning from prose. <u>Journal of Educational Psychology</u>, 1971, 62, 387-394.
- Wickens, D. D. Encoding categories of words: An empirical approach to meaning. Psychological Review, 1970, 77, 1-15.

# DATE